

CONVERGENCE RESULTS FOR PROJECTED LINE-SEARCH METHODS ON VARIETIES OF LOW-RANK MATRICES VIA ŁOJASIEWICZ INEQUALITY

REINHOLD SCHNEIDER[†] AND ANDRÉ USCHMAJEV[‡]

Abstract. As an initial step towards low-rank optimization algorithms using hierarchical tensors, the aim of this paper is to derive convergence results for projected line-search methods on the real-algebraic variety $\mathcal{M}_{\leq k}$ of real $m \times n$ matrices of rank at most k . Such methods extend successfully used Riemannian optimization methods on the smooth manifold \mathcal{M}_k of rank- k matrices to its closure by taking steps along gradient-related directions in the tangent cone, and afterwards projecting back to $\mathcal{M}_{\leq k}$. Considering such a method circumvents the difficulties which arise from the non-closedness and the unbounded curvature of \mathcal{M}_k . The point-wise convergence is obtained for real-analytic functions on the basis of a Łojasiewicz inequality for the projection of the negative gradient to the tangent cone. If the derived limit point lies on the smooth part of $\mathcal{M}_{\leq k}$, i.e. in \mathcal{M}_k , this boils down to more or less known results, but with the benefit that asymptotic convergence rate estimates (for specific step-sizes) can be obtained without an a-priori curvature bound, simply from the fact that the limit lies on a smooth manifold. At the same time, one can give a convincing justification for assuming critical points to lie in \mathcal{M}_k : if X is a critical point of f on $\mathcal{M}_{\leq k}$, then either X has rank k , or $\nabla f(X) = 0$.

Key words. Convergence analysis, line-search methods, low-rank matrices, Riemannian optimization, steepest descent, Łojasiewicz gradient inequality, tangent cone

AMS subject classifications. 65K06, 40A05, 26E05, 65F30, 15B99, 15A83,

1. Introduction. This paper is concerned with line-search algorithms for low-rank matrix optimization. Let $k \leq \min(m, n)$. The set

$$\mathcal{M}_k = \{X \in \mathbb{R}^{m \times n} : \text{rank}(X) = k\}$$

of real rank- k matrices is a real-analytic submanifold of $\mathbb{R}^{m \times n}$. Thus, in order to approach a solution of

$$\min_{X \in \mathcal{M}_k} f(X), \tag{1.1}$$

where $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ is continuously differentiable, one can use the algorithms known from Riemannian optimization, the simplest being the projected gradient method (or steepest descent method)

$$X_{n+1} = R(X_n + \alpha_n \Pi_{T_{X_n} \mathcal{M}_k}(-\nabla f(X_n))). \tag{1.2}$$

Here, $\Pi_{T_{X_n} \mathcal{M}_k}$ is the orthogonal projection on the tangent space at X_n , $\alpha_n \geq 0$ is a step-size, and R is a *retraction*, which takes vectors from the affine tangent plane back to the manifold [2, 40]. Riemannian optimization on \mathcal{M}_k (and other matrix manifolds) has become an important tool for low-rank approximation in several applications, e.g. solutions of matrix equations such as the Lyapunov equation, model reduction in machine learning, low-rank matrix completion, and others, see, for instance, [12, 31, 32, 39, 43, 42]. Typically, more sophisticated methods than steepest descent, like nonlinear conjugate gradients, Newton's method, or line-search along geodesics, are employed. In any case, the search directions in a line-search method need to be sufficiently gradient-related in order to enabling for a convergence analysis.

An alternative interpretation of the projected gradient method (1.2) is that of a discretized gradient flow satisfying the Dirac–Frenkel variational principle, i.e., of the integration of the ODE

$$\dot{X}(t) = \Pi_{T_{X(t)} \mathcal{M}_k}(-\nabla f(X(t)))$$

using Euler's explicit method with some step-size strategy. Therefore, our studies are closely related to the growing field of dynamical low-rank approximation of ODEs [17, 35, 29], especially when they admit a strict Lyapunov function.

[†]Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany. E-mail: schneider@math.tu-berlin.de

[‡]MATHICSE-ANCHP, Section de Mathématiques, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland. E-mail: andre.uschmajew@epfl.ch

The convergence analysis of sequences in \mathcal{M}_k is hampered by the fact that this manifold is not closed in the ambient space $\mathbb{R}^{m \times n}$. The manifold properties break down at the boundary which consists of matrices of rank less than k . It might happen that a minimizing sequence for (1.1) needs to cross such a singular point, or even converge to it, for instance when the rank of the final solution has been overestimated. Also, the formal domain of definition of a smooth retraction becomes tiny at points close to singularities, leading to too small allowed step-sizes in theory. Even if these objections pose no serious problems in practice, they make it difficult to derive a-priori convergence statements without making unjustified assumptions on the smallest singular values or adding regularization, cf. [20, 17, 18, 28, 42].

It certainly would be more convenient to optimize and analyze on the closure (stratification)

$$\mathcal{M}_{\leq k} = \{X \in \mathbb{R}^{m \times n} : \text{rank}(X) \leq k\} = \bigcup_{0 \leq s \leq k} \mathcal{M}_s, \quad (1.3)$$

which is a real-algebraic variety. Indeed, in many applications one is fine with any solution of the problem

$$\min_{X \in \mathcal{M}_{\leq k}} f(X). \quad (1.4)$$

There is in principle no difficulty to devise line-search methods on $\mathcal{M}_{\leq k}$. First, in singular points, one has to use search directions in the tangent cone (instead of tangent space), for instance a projection of the negative gradient on the tangent cone. The tangent cones of $\mathcal{M}_{\leq k}$ are explicitly known [12], and projecting on them is easy (see Theorem 3.2 and Corollary 3.3). Second, one needs a “retraction” that maps from the affine tangent cone back to $\mathcal{M}_{\leq k}$, a most natural choice being the Euclidean projection

$$R(X_n + \Xi) = \operatorname{argmin}_{Y \in \mathcal{M}_{\leq k}} \|Y - (X_n + \Xi)\|_F, \quad (1.5)$$

which can be calculated using singular value decomposition. The aim of this paper is to develop convergence results for such a method based on a Łojasiewicz inequality for the projected negative gradient.

The convergence theory of gradient flows based on the Łojasiewicz gradient inequality [27] has, together with results based on a more general tool, the Łojasiewicz-Kurdyka inequality [21, 8, 9], attracted much attention in nonlinear optimization during the last years [1, 4, 5, 6, 7, 11, 22, 23, 24, 25, 26, 30, 34, 44]. In part, this interest seems to have been triggered by the paper [1], where the following theorem was proved.

THEOREM. *Let $f : \mathbb{R}^N \rightarrow \mathbb{R}$ be continuously differentiable, and let $(x_n) \subseteq \mathbb{R}^N$ be a sequence of iterates satisfying the strong descent conditions*

$$\begin{aligned} f(x_{n+1}) - f(x_n) &\leq -\sigma \|\nabla f(x_n)\| \|x_{n+1} - x_n\| \quad (\text{for some } \sigma > 0), \\ f(x_{n+1}) &= f(x_n) \quad \Rightarrow \quad x_{n+1} = x_n. \end{aligned} \quad (1.6)$$

Assume further that the sequence possesses a cluster point x^ that satisfies the Łojasiewicz gradient inequality, i.e., there exist $\theta > 0$ and $\Lambda > 0$ such that*

$$|f(y) - f(x^*)|^{1-\theta} \leq \Lambda \|\nabla f(y)\| \quad (1.7)$$

for all y in some neighborhood of x^ . Then x^* is the limit of the sequence (x_n) .*

For instance, real-analytic functions satisfy the Łojasiewicz gradient inequality in every point. It is possible to obtain a stronger results if a small step-size safe-guard of the form

$$\|x_{n+1} - x_n\| \geq \kappa \|\nabla f(x_n)\| \quad (\text{for some } \kappa > 0) \quad (1.8)$$

can be assumed. Not only can one then conclude that the limit x^* is a critical point of f . Also, the asymptotic convergence rate in terms of the Łojasiewicz parameters θ and Λ can be estimated along lines developed, e.g., in [4, 7, 25, 30]. No second order information is required, but a linear convergence rate can only be established

when $\theta = 1/2$, which in general cannot be checked in advance. These type of results can be applied to classical line-search algorithms in \mathbb{R}^N when minimizing a real-analytic function using an angle condition for the search directions and Armijo back-tracking for the step-size selection [1].

The theory can be generalized to gradient systems on real-analytic manifolds. Lageman [23] considered descent iterations on Riemannian manifolds via families of local parametrizations, with retracted line-search methods like (1.2) being a special case of his setting. Convergence results are obtained by making regularity assumptions on the used family of parametrizations. Merlet and Nguyen [30] consider a discrete projected θ -scheme for integrating an ODE on a smooth embedded submanifold. They prove the existence of step-sizes which ensure convergence (with some rates) via Łojasiewicz inequality by assuming a uniform bound on the second order terms in the Taylor expansion of the metric projection. The main problem with the non-compact submanifold \mathcal{M}_k , without which the following considerations would be unnecessary, is that such an assumption is unjustified. The second order term in the metric projection scales like the inverse of the smallest singular value [17, 3], which gets arbitrarily large in case the iterates would approach the boundary, even if they remained bounded. It seems, however, that such a uniform bound for the projection is not needed if one is willing to sacrifice the exact knowledge of the constant in the convergence rate estimates: if a gradient projection method (x_n) on a smooth manifold converges to *some* point of it, one will have *some* curvature bound in the vicinity of the limit, implying a minimum step-size safe-guard like (1.8) for *large enough* n .

Therefore, our plan is this: Via a version of the Łojasiewicz inequality for projections of the negative gradients on tangent cones, we deduce that the iterates of a line-search method on $\mathcal{M}_{\leq k}$ with a particular choice of step-sizes (Armijo) do converge. This would not be possible on \mathcal{M}_k . Then we assume that the limit has full rank, i.e., lies in \mathcal{M}_k , to establish that it is a critical point, and to estimate the convergence rate. This seems circular at first sight, but it is not: operating on $\mathcal{M}_{\leq k}$ allows to make arbitrary steps in the tangent cone without any formal restrictions. On the other hand, the full rank assumption on the limit is very natural, due to the following, probably new insight (repeated as Corollary 3.4).

THEOREM. *Let $k \leq \min(m, n)$, and let $X^* \in \mathcal{M}_{\leq k}$ be a critical point of (1.4) (see section 2.1). Then either $\text{rank}(X^*) = k$ or $\nabla f(X^*) = 0$.*

Accordingly, it typically will be impossible to prove convergence to a rank-deficient critical point by a method which (in regular points) only “sees” projections of the gradient on lower-dimensional subspaces.

Contributions and outline. The paper has two parts: in section 2 abstract convergence statements for line-search methods are established. In section 3 these are applied to line-search methods on $\mathcal{M}_{\leq k}$. We focus on real-analytic objective functions f that are bounded below. The following list highlights the results.

- In section 2.2 we state the abstract conditions to obtain convergence and convergence rate estimates for descent methods on a closed set $\mathcal{M} \subseteq \mathbb{R}^N$ via Łojasiewicz inequalities for projections of the negative gradient on the tangent cones (Theorem 2.2). This is probably not the most general way, but is sufficiently simple for our later needs. Up to the replacement of the gradient by this projected gradient, the arguments are well-known.
- In section 2.3 we define line-search schemes using gradient-related search directions on tangent cones (Algorithm 1). The step-sizes are selected by an Armijo rule, and can hence be found by back-tracking. The used generalized retractions still need to be second-order approximations of the identity in every fixed tangent direction (possibly non-uniform w.r.t the direction), see Definition 2.3. This definition is tailored to tangential projections on algebraic varieties. The main convergence result is Corollary 2.10: If f is real-analytic, then a cluster point of the sequence generated by the method, in whose neighborhood \mathcal{M} forms a real-analytic submanifold, must be its limit, a critical point of the problem, and the asymptotic convergence rate can be estimated. These estimates contain unknown constants, but in turn no convexity, and no curvature bounds are required to obtain them.
- Section 3.1 is devoted to the tangent cones of $\mathcal{M}_{\leq k}$. We give a much shorter derivation of their structure (Theorem 3.2) compared to [12]. The projection on the tangent cone is a simple and feasible operation (cf. Algorithm 2). Astonishingly, when $\text{rank}(X) < k$ (i.e. when the tangent cone

is not just a linear space), the norm of this projection can be estimated from below by the norm of the projected matrix itself, independently of X (Corollary 3.3). This implies the above a-priori statement on the rank of critical points, and also immediately proves a Łojasiewicz inequality for the projected negative gradient whenever it holds for the full gradient (cf. Theorem 3.7), an observation which lets us expect that subsequent results can be easily extended to larger classes of functions than real-analytic ones.

- Finally, in sections 3.3 and 3.4 we consider two concrete line-search methods on $\mathcal{M}_{\leq k}$, the first one being the classical steepest descent method with projection (Algorithm 3), the second using search directions which do not increase the rank, making it an elegant, retraction-free method (Algorithm 4). When f is real-analytic, the strongest convergence results for these methods are obtained under the assumption that regular cluster points in \mathcal{M}_k exist (Theorems 3.8 and 3.9). We compare both algorithms for a toy example of matrix completion.

In general, the results are restricted to finite-dimensional spaces, and one has to expect that hidden constants, and the provable convergence rate itself (in case of linear convergence), deteriorate with the problem size. This is to be expected from a black-box tool like the Łojasiewicz inequality which cannot be easily extended to infinite dimension, cf. [15, 16]. The limitation to finite dimension has been disregarded in related works as well [11, 25]. On the other hand, Vandereycken [42] for instance observed more or less dimension-independent convergence rates for matrix completion of synthetic data using nonlinear CG.

We refrained from including serious numerical experiments, as they can be found elsewhere, and do not provide much insight into our theoretical results: linear convergence is nearly always observed, but the rates in our theorems are generally not explicit. Another point is that it is difficult to design a method that will have rank-deficient iterates except at a prescribed starting point. In practice, rank-deficiency in the limiting stage of a numerical iteration are rarely observed, which is somehow explained by the theorem stated above. We therefore repeat our motivation for considering a method on $\mathcal{M}_{\leq k}$ at all: in a method on \mathcal{M}_k we have a formal limitation of the step size in the tangent space to prevent rank-deficiency, which is directly concerned with the distance of the iterates from singular points. Often, there is no argument to guarantee that this distance will remain bounded from below. The idea that has driven this work is that this rather theoretical issue can be elegantly circumvented: a line-search method on $\mathcal{M}_{\leq k}$ can deal with potential singular points in theory, although in the most likely cases it will not encounter one in a real computation. Thus, in the end it will not differ from a line-search method on \mathcal{M}_k as used in practice, thereby establishing its convergence.

2. Convergence of gradient methods via Łojasiewicz inequality. Let $\mathcal{D} \subseteq \mathbb{R}^N$ be open, and $f : \mathcal{D} \rightarrow \mathbb{R}$. In the whole paper, unless something else is stated, we assume:

$$f \text{ is continuously differentiable and bounded below.} \quad (\text{A0})$$

Together with f we consider the minimization problem

$$\min_{x \in \mathcal{M}} f(x) \quad (2.1)$$

on a *closed* subset $\mathcal{M} \subset \mathcal{D}$ and assume it to have a solution.

By $\|\cdot\|$ we denote the usual Euclidean norm on \mathbb{R}^N . The natural numbers \mathbb{N} contain zero.

2.1. Optimality condition. Let $x \in \mathcal{M}$. The *tangent cone* (also called *contingent cone*) at x is

$$T_x \mathcal{M} = \{\xi \in \mathbb{R}^N : \exists (x_n) \subseteq \mathcal{M}, (a_n) \subseteq \mathbb{R}^+ \text{ s.t. } a_n(x_n - x) \rightarrow \xi\}, \quad (2.2)$$

see, e.g., [13, 38]. It is a closed cone. Since it is in general not convex, a projection onto $T_x \mathcal{M}$ may not be uniquely defined. However, let $y \in \mathbb{R}^N$, then any $z \in T_x \mathcal{M}$ with $\|y - z\| = \text{dist}_{\|\cdot\|}(y, T_x \mathcal{M})$ is orthogonal on the error, in the sense that

$$\|z\|^2 = \|y\|^2 - \|y - z\|^2 = \|y\|^2 - \text{dist}_{\|\cdot\|}(y, T_x \mathcal{M})^2.$$

Therefore, we define

$$p_x : \mathbb{R}^N \rightarrow [0, \infty), \quad p_x(y) = \sqrt{\|y\|^2 - \text{dist}_{\|\cdot\|}(y, T_x \mathcal{M})^2},$$

and

$$g^-(x) = p_x(-\nabla f(x)).$$

It is the norm of any “orthogonal” projection of $-\nabla f(x)$ onto $T_x \mathcal{M}$. A characterization, which resembles the norm of a restricted linear operator, is

$$g^-(x) = - \max_{\substack{\xi \in T_x \mathcal{M} \\ \|\xi\| \leq 1}} \nabla f(x)^\top \xi, \quad (2.3)$$

and the maximum is achieved if and only if $g^-(x)\xi$ is a best approximation of $-\nabla f(x)$ in $T_x \mathcal{M}$.

The *polar tangent cone*

$$T_x^\circ \mathcal{M} = \{y \in \mathbb{R}^N : y^\top \xi \leq 0 \text{ for all } \xi \in T_x \mathcal{M}\} \quad (2.4)$$

is always a closed convex cone. The necessary first-order optimality condition for x^* to be a relative local minimum of f on \mathcal{M} is

$$-\nabla f(x^*) \in T_{x^*}^\circ \mathcal{M},$$

see again, e.g., [13, 38]. By (2.3), the statements $y \in T_x^\circ \mathcal{M}$ and $p_x(y) = 0$ are equivalent. Therefore, the equivalent first-order optimality condition we work with in this paper is

$$g^-(x^*) = 0.$$

In the case that $T_x \mathcal{M}$ is a linear space, $T_x^\circ \mathcal{M}$ is its orthogonal complement, $g^-(x)$ is the norm of the orthogonal projection of $\nabla f(x)$, and everything that has been said becomes quite evident. Moreover, if \mathcal{M} is a differentiable manifold and ∇f continuous, then g^- is continuous on \mathcal{M} . In general, it is not.

2.2. General convergence theorem. To solve (2.1), we consider an iteration $(x_n) \subseteq \mathcal{M}$. Throughout the paper we will use the shorthands

$$f_n = f(x_n), \quad \nabla f_n = \nabla f(x_n), \quad g_n^- = g^-(x_n), \quad \text{and} \quad T_n \mathcal{M} = T_{x_n} \mathcal{M}.$$

We make the following assumptions.

- **Primary descent condition:** There exists $\sigma > 0$ such that for large enough n it holds

$$f_{n+1} - f_n \leq -\sigma g_n^- \|x_{n+1} - x_n\|. \quad (\mathbf{A1})$$

- **Stationary condition:** For large enough n it holds

$$g_n^- = 0 \quad \Rightarrow \quad x_{n+1} = x_n. \quad (\mathbf{A2})$$

- **Łojasiewicz inequality for projected negative gradient:** For every $x \in \mathcal{M}$ with $g^-(x) = 0$ there exist $\delta > 0$, $\Lambda > 0$, and $\theta \in (0, 1/2]$ such that for all $y \in \mathcal{M}$ with $\|y - x\| < \delta$ it holds

$$|f(y) - f(x)|^{1-\theta} \leq \Lambda g^-(y). \quad (\mathbf{A3})$$

- **Asymptotic small step-size safe-guard:** There exists $\kappa > 0$ such that for large enough n it holds

$$\|x_{n+1} - x_n\| \geq \kappa g_n^-. \quad (\mathbf{A4})$$

The Łojasiewicz inequality holds for instance for real-analytic functions in open domains of finite-dimensional linear spaces [27]. Using local real-analytic charts, it is an easy task to prove a version for real-analytic manifolds. (Recall that $g^-(x)$ then is just the norm of the orthogonal projection of $\nabla f(x)$ on the tangent space.) We note this for later reference.

PROPOSITION 2.1. *Let f be real-analytic, and $\mathcal{M} \subseteq D$ be a real-analytic submanifold of \mathbb{R}^N . Then **(A3)** holds at any point $x \in \mathcal{M}$ (and not just at points with $g^-(x) = 0$).*

In section 3.3 we prove the Łojasiewicz inequality for the set $\mathcal{M}_{\leq k}$.

We have the following general convergence theorem.

THEOREM 2.2. *Under assumptions **(A1)**–**(A3)**, if there exists a cluster point x^* of the sequence (x_n) , it is actually its limit point. If further **(A4)** holds, then the convergence rate can be estimated by*

$$\|x_n - x^*\| \lesssim \begin{cases} e^{-cn} & \text{if } \theta = \frac{1}{2} \text{ (for some } c > 0), \\ n^{-\frac{\theta}{1-2\theta}} & \text{if } 0 < \theta < \frac{1}{2}. \end{cases}$$

Moreover, $g_n^- \rightarrow 0$.

The question when $g_n^- \rightarrow 0$ actually does imply $g^-(x^*) = 0$ is delicate. A sufficient condition would be $T_{x^*}\mathcal{M} \subset \liminf_{n \rightarrow \infty} T_n\mathcal{M}$ in the sense of set convergence (see, e.g., [38]). Unfortunately, this will usually not hold in the singular points of $\mathcal{M}_{\leq k}$ when approached by a sequence of full-rank matrices (Theorem 3.2). We later will be forced to make some smoothness assumptions on a neighborhood of x^* .

The above theorem is an adaption of similar results scattered through the literature. Up to replacing the usual gradient by the projected negative gradient, assumptions **(A1)**, **(A2)** and **(A3)** are the same as in [1, Theorem 3.2], and are sufficient to prove the convergence of the sequence (x_n) if it is bounded. **(A2)** is a natural technical requirement to the algorithm to not move in the critical point set, and typically satisfied if the iteration is gradient-related. Adding assumption **(A4)** does not only guarantee that the limit point is critical, but allows to estimate the convergence rate along known lines, e.g. [4, 25]. However, as **(A4)** is required here only for n larger than some unknown n_0 , one cannot determine the constants in the convergence rate estimate explicitly (a constant depending on n_0 can be deduced from the proof below). Results for manifolds can be found in [23, 24, 30]. In this context, we should remark that the ambient norm $\|x_{n+1} - x_n\|$, as used in **(A1)** and **(A4)**, is not necessarily a natural measure of distance on \mathcal{M} , but is particularly appropriate when the restriction to \mathcal{M} is motivated to reduce the complexity of a minimization problem in \mathbb{R}^N , as it typically is the case for low-rank optimization. Although no changes in the known arguments besides replacing $\|\nabla f\|$ by g^- are required, we give a proof of Theorem 2.2 to keep the paper self-contained.

Proof of Theorem 2.2. We can assume that $g_n^- > 0$ for all n since otherwise the sequence becomes stationary and there is nothing to prove. There will also be no loss in generality to assume that **(A1)** and **(A2)** hold for all n , and that $f(x^*) = 0$. Then $0 \leq f(x^*) \leq f_n$ for all n , and the Łojasiewicz inequality at x^* reads

$$f(x)^{1-\theta} \leq \Lambda g^-(x) \tag{2.5}$$

whenever $\|x - x^*\| < \delta = \delta(x^*)$. Let $\epsilon \in (0, \delta]$, and assume $\|x_n - x^*\| < \delta$. Then, by (2.5) and **(A1)**,

$$\|x_n - x_{n+1}\| \leq \frac{\Lambda}{\sigma} f_n^{\theta-1} (f_n - f_{n+1}).$$

Using that for $\varphi \in [f_{n+1}, f_n]$ there holds $f_n^{\theta-1} \leq \varphi^{\theta-1} \leq f_{n+1}^{\theta-1}$, we can estimate

$$f_n^{\theta-1} (f_n - f_{n+1}) \leq \int_{f_{n+1}}^{f_n} \varphi^{\theta-1} d\varphi = \frac{1}{\theta} (f_n^\theta - f_{n+1}^\theta),$$

and thus obtain

$$\|x_n - x_{n+1}\| \leq \frac{\Lambda}{\sigma\theta} (f_n^\theta - f_{n+1}^\theta).$$

More generally, let $\|x_k - x^*\| < \epsilon \leq \delta$ for $n \leq k < m$, we get by this argument that

$$\|x_m - x_n\| \leq \sum_{k=n}^m \|x_{k+1} - x_k\| \leq \sum_{k=n}^m \frac{\Lambda}{\sigma\theta} (f_k^\theta - f_{k+1}^\theta) = \frac{\Lambda}{\sigma\theta} (f_n^\theta - f_m^\theta) \leq \frac{\Lambda}{\sigma\theta} f_n^\theta. \quad (2.6)$$

Since x^* is an accumulation point, we can pick n so large that

$$\|x_n - x^*\| < \frac{\epsilon}{2} \quad \text{and} \quad \frac{\Lambda}{\sigma\theta} f_n^\theta < \frac{\epsilon}{2}.$$

Then (2.6) inductively implies $\|x_m - x^*\| < \epsilon$ for all $m \geq n$. This proves that x^* is the limit point of the sequence, and, by (A4), $g_n^- \rightarrow 0$.

To estimate the convergence rate, let $r_n = \sum_{k=n}^\infty \|x_{k+1} - x_k\|$. Then $\|x_n - x^*\| \leq r_n$, so it suffices to estimate the latter. By (2.6), (2.5), and (A4), there exists $n_0 \geq 1$ such that for $n \geq n_0$ it holds that

$$r_n^{\frac{1-\theta}{\theta}} \leq \left(\frac{\Lambda}{\sigma\theta}\right)^{\frac{1-\theta}{\theta}} f_n^{1-\theta} \leq \left(\frac{\Lambda}{\sigma\theta}\right)^{\frac{1-\theta}{\theta}} \frac{\Lambda}{\kappa} \|x_{n+1} - x_n\| = \left(\frac{\Lambda}{\sigma\theta}\right)^{\frac{1-\theta}{\theta}} \frac{\Lambda}{\kappa} (r_n - r_{n+1}),$$

that is,

$$r_{n+1} \leq r_n - \nu r_n^{\frac{1-\theta}{\theta}} \quad (2.7)$$

with $\nu = \left(\frac{\Lambda}{\sigma\theta}\right)^{\frac{\theta-1}{\theta}} \frac{\kappa}{\Lambda}$. Now, if $\theta = 1/2$, we get from (2.7) that $\nu \in (0, 1)$, and

$$r_n \leq r_{n_0} (1 - \nu)^{n-n_0} \left(e^{\ln(1-\nu)}\right)^n$$

for $n \geq n_0$. The case $0 < \theta < 1/2$ is more delicate. We follow Levitt [25]: put $p = \frac{\theta}{1-2\theta}$, $C \geq \max((\frac{\nu}{p})^{-p}, r_{n_0} n_0^{-p})$, and $s_n = C n^{-p}$, then $s_{n_0} \geq r_{n_0}$, and

$$s_{n+1} = s_n (1 + n^{-1})^{-p} \geq s_n (1 - p n^{-1}) = s_n - \frac{p}{C^{1/p}} s_n^{\frac{p+1}{p}} \geq s_n - \nu s_n^{\frac{p+1}{p}} = s_n - \nu s_n^{\frac{1-\theta}{\theta}}$$

(the first inequality holding by convexity of x^{-p}). Using induction, it now follows from (2.7) that $r_n \leq s_n$ for all $n \geq n_0$. \square

2.3. Relation to retracted line-search methods. For line-search methods in \mathbb{R}^N it is well known how to obtain convergence results based on the Łojasiewicz inequality [1]. Here we provide results for projected gradient flows on a set \mathcal{M} . In the introduction we pointed to earlier results in this direction.

2.3.1. Retractions. Following [2], a retracted line-search method on a smooth manifold \mathcal{M} has the general form

$$x_0 \in \mathcal{M}, \quad x_{n+1} = R(x_n + \alpha_n \xi_n), \quad (2.8)$$

where ξ_n are tangent vectors at x_n , $\alpha_n \geq 0$, and $R : T\mathcal{M} \rightarrow \mathcal{M}$ is a *smooth retraction* [40].¹ That is, R is a C^∞ map which takes vectors $x + \xi_x$ from the tangent plane at x back to the manifold, and has the additional property of being a first-order approximation of the exponential map, which means

$$\lim_{T_x \mathcal{M} \ni \xi_x \rightarrow 0} \frac{\|R(x + \xi_x) - (x + \xi_x)\|}{\|\xi_x\|} = 0 \quad (2.9)$$

¹We identified $x + \xi_x$ with the element (x, ξ_x) from the tangent bundle $T\mathcal{M}$.

for all $x \in \mathcal{M}$. In other words, the first derivative of the map $\xi \mapsto R(x + \xi)$ in 0 should be the identity (on $T_x\mathcal{M}$). However, since we do not want to restrict ourselves to smooth manifolds, we make the following, slightly more general definition.

DEFINITION 2.3 (Retraction). *Let \mathcal{M} be closed. A map*

$$R : \bigcup_{x \in \mathcal{M}} \{x\} \times T_x\mathcal{M} \rightarrow \mathcal{M}$$

(where now $T_x\mathcal{M}$ is the tangent cone) will be called a retraction, if for any fixed $x \in \mathcal{M}$ and $\xi_x \in T_x\mathcal{M}$ it holds that $\alpha \mapsto R(x + \alpha\xi_x)$ is continuous on $[0, \infty)$, and

$$\lim_{\alpha \rightarrow 0^+} \frac{R(x + \alpha\xi_x) - (x + \alpha\xi_x)}{\alpha} = 0. \quad (2.10)$$

The existence of such a retraction has implications on the regularity of the set \mathcal{M} . It is equivalent to the (one sided) differentiability of the distance map $\alpha \mapsto \text{dist}_{\|\cdot\|}(x + \alpha\xi_x, \mathcal{M})$ along tangent rays. This is for instance the case for real-algebraic varieties like $\mathcal{M}_{\leq k}$, which follows from the fact that for every tangent vector ξ_x to an algebraic variety there exists an analytic arc $\gamma : [0, \epsilon) \rightarrow \mathcal{M}$ such that $\xi_x = \dot{\gamma}(0)$ [37, Proposition 2].

By (2.10), $R(x + \alpha\xi_x)$ is a good approximation of $x + \alpha\xi_x$ for very small $\alpha > 0$. In particular, for any fixed ξ_x and $\epsilon > 0$ we have

$$(1 - \epsilon)\alpha\|\xi_x\| \leq \|R(x + \alpha\xi_x) - x\| \leq (1 + \epsilon)\alpha\|\xi_x\| \quad (2.11)$$

for sufficiently small α . It means that a (small enough) step made in the tangent cone is neither increased nor decreased too much by the retraction, which obviously is of importance to analyze a line-search method like (2.8). For what follows, we assume that we have a general upper bound for arbitrary steps:

$$\|R(x + \xi_x) - x\| \leq M\|\xi_x\| \quad \text{for all } x \in \mathcal{M} \text{ and } \xi_x \in T_x\mathcal{M}. \quad (2.12)$$

This imposes no serious restriction.

Since \mathcal{M} is assumed to be closed, the natural choice for R , though practically not always the most convenient, is the the best approximation of $x + \xi_x$ in the Euclidean ambient norm (metric projection), that is,

$$R(x + \xi_x) \in \underset{y \in \mathcal{M}}{\text{argmin}} \|y - (x + \xi_x)\|. \quad (2.13)$$

By the remarks above, this defines a valid retraction for example on closed real-algebraic varieties (cf. (3.9)) with $M = 2$ in (2.12). For the variety $\mathcal{M}_{\leq k}$ of bounded rank matrices one even can take $M = 1 + 2^{-1/2}$ (Proposition 3.6).

2.3.2. Angle condition. To obtain such strong convergence results as we have in mind, one naturally has to guarantee that the search direction ξ_n in (2.8) remain sufficiently gradient-related. We call $\xi_n \in T_n\mathcal{M}$ a *descent direction*, if $\nabla f_n^T \xi_n < 0$.

DEFINITION 2.4 (Angle condition). *Given $x_n \in \mathcal{M}$ and $\omega \in (0, 1]$, $\xi_n \in T_n\mathcal{M}$ is said to satisfy the ω -angle condition, if*

$$\nabla f_n^T \xi_n \leq -\omega g_n^- \|\xi_n\|. \quad (2.14)$$

An equivalent statement is that the overlap between $-\nabla f_n / \|\nabla f_n\|$ and $\xi_n / \|\xi_n\|$ is at least $\omega g_n^- / \|\nabla f_n\|$.

For clarity, we emphasize the following.

PROPOSITION 2.5. *Any Euclidean best-approximation $\xi_n \in \underset{\xi \in T_n\mathcal{M}}{\text{argmin}} \|\xi - \nabla f_n\|$ of $-\nabla f_n$ on $T_n\mathcal{M}$ satisfies the angle condition with $\omega = 1$. Moreover, with this choice, $\xi_n = 0$ if and only if $g_n^- = 0$.*

Proof. As discussed in section 2.1, it holds in this case that $g_n^- = \|\xi_n\| = \sqrt{\|\nabla f_n\|^2 - \|\nabla f_n + \xi_n\|^2}$, which implies $\nabla f_n^T \xi_n = -g_n^- \|\xi_n\|$. \square

2.3.3. Armijo point. Given $x_n \in \mathcal{M}$, and a descent direction $\xi_n \in T_n\mathcal{M}$, we will have to pick a step-size α_n small enough to satisfy **(A1)**. It should, however, be as large as possible in order to hopefully guarantee **(A4)**.

DEFINITION 2.6 (Armijo point). *Let $\xi_n \in T_n\mathcal{M}$ be a descent direction at $x_n \in \mathcal{M}$, $\bar{\beta}_n > 0$, and $\beta, c \in (0, 1)$. The number*

$$\alpha_n = \max(\{\beta^m \bar{\beta}_n : m \in \mathbb{N}, f(R(x_n + \beta^m \bar{\beta}_n \xi_n)) - f_n \leq c \beta^m \bar{\beta}_n \nabla f_n^\top \xi_n\} \cup \{0\})$$

is called the Armijo point for $x_n, \xi_n, \bar{\beta}_n, \beta, c$.

This will be our choice for the step-size α_n in all subsequent algorithms. The importance of the Armijo point lies in the fact that it in principle it can be found in finitely many steps using back-tracking, if it is positive. Concerning the question when $\alpha_n > 0$ holds, we introduce, as usual, another important point:

$$\bar{\alpha}_n = \min\{\alpha > 0 : f(R(x_n + \alpha \xi_n)) - f_n = c \alpha \nabla f_n^\top \xi_n\}. \quad (2.15)$$

PROPOSITION 2.7. *Assume **(A0)**. Let $\xi_n \in T_n\mathcal{M}$ be a descent direction at $x_n \in \mathcal{M}$, and $\beta, c \in (0, 1)$. Then $\bar{\alpha}_n > 0$ exists, and $f(R(x_n + \alpha \xi_n)) - f_n \leq c \alpha \nabla f_n^\top \xi_n$ for all $\alpha \in [0, \bar{\alpha}_n]$. The Armijo point α_n satisfies*

$$\begin{aligned} \alpha_n &\geq \beta \bar{\alpha}_n && \text{if } \bar{\beta}_n > \bar{\alpha}_n, \\ \alpha_n &= \bar{\beta}_n && \text{if } \bar{\beta}_n \leq \bar{\alpha}_n. \end{aligned}$$

Proof. For convenience, let $\hat{R}(\alpha) = R(x_n + \alpha \xi_n)$ and $F(\alpha) = f_n + c \alpha \nabla f_n^\top \xi_n$. We need to show, that $f(\hat{R}(\alpha))$ is strictly smaller than $F(\alpha)$ for sufficiently small $\alpha > 0$. By Taylor's theorem and (2.10),

$$\begin{aligned} f(\hat{R}(\alpha)) &= f_n + \nabla f_n^\top (\hat{R}(\alpha) - x_n) + o(\|\hat{R}(\alpha) - x_n\|) \\ &= f_n + \nabla f_n^\top (\alpha \xi_n + o(\alpha)) + o(\|\hat{R}(\alpha) - x_n\|) \\ &= f_n + c \alpha \nabla f_n^\top \xi_n + (1 - c) \alpha \nabla f_n^\top \xi_n + o(\alpha) + o(\|\hat{R}(\alpha) - x_n\|), \end{aligned}$$

where $o(h)$ denotes a quantity with $o(h)/h = 0$ for $h \rightarrow 0^+$. By (2.11), the ratio $\|\hat{R}(\alpha) - x_n\|/\alpha$ converges to $\|\xi_n\|$ for $\alpha \rightarrow 0^+$, which implies $o(\|\hat{R}(\alpha) - x_n\|) = o(\alpha)$. As desired, it now follows that

$$\frac{1}{\alpha} (f(\hat{R}(\alpha)) - f_n - c \alpha \nabla f_n^\top \xi_n) = (1 - c) \nabla f_n^\top \xi_n + \frac{o(\alpha)}{\alpha}$$

is negative for small enough α . Since $\alpha \mapsto f(R(x_n + \alpha \xi_n))$ is continuous and bounded below by **(A0)**, while $F(\alpha)$ is not, the smallest positive intersection point $\bar{\alpha}_n$ must exist. The assertions on α_n are immediate. \square

The role of the parameter $\bar{\beta}_n$ in Definition 2.6 is to adjust the initial length of the search direction ξ_n on which no assumptions have been made. When $\bar{\beta}_n \|\xi_n\|$ is too small one has no chance to establish a minimum step-size safe-guard like **(A4)**. The restriction we make is

$$\bar{\beta}_n \geq \min \left(\frac{g_n^-}{\|\xi_n\|}, \bar{\alpha}_n \right). \quad (2.16)$$

Note that one can achieve $\bar{\beta}_n > \bar{\alpha}_n$ in finitely many steps, but how many may depend on n .

2.3.4. Convergence results. The algorithm we analyze is formalized as Algorithm 1. By Propositions 2.5 and 2.7, all steps are feasible. The mere convergence of the produced sequence (x_n) under a Lojasiewicz assumptions is guaranteed by Theorem 2.2.

COROLLARY 2.8. *Assume **(A0)**. The sequence (x_n) produced by Algorithm 1 satisfies **(A1)** with $\sigma = \omega c M^{-1}$ (M being the constant from (2.12)), and **(A2)**. Consequently, if a cluster point x^* exists and satisfies the Lojasiewicz inequality **(A3)**, then $\lim_{n \rightarrow \infty} x_n = x^*$.*

Algorithm 1: Gradient-related projection method with line-search

Input: Starting point $x_0 \in \mathcal{M}$, $\omega \in (0, 1]$, $\beta, c \in (0, 1)$.

- 1 **for** $n=0, 1, 2, \dots$ **do**
- 2 Choose $\xi_n \in T_n \mathcal{M}$ satisfying the ω -angle condition, but choose $\xi_n = 0$ only when $g_n^- = 0$;
- 3 Choose $\bar{\beta}_n \geq \min(g_n^- / \|\xi_n\|, \bar{\alpha}_n)$, and find Armijo point α_n for $x_n, \xi_n, \bar{\beta}_n, \beta, c$;
- 4 Form the next iterate

$$x_{n+1} = R(x_n + \alpha_n \xi_n).$$
- 5 **end**

Proof. Property **(A1)** follows immediately from (2.14) and (2.12), **(A2)** holds by construction. \square

Obviously, it is not necessary to choose the Armijo step-size to obtain this result, it suffices to have $f(R(x_n + \alpha_n \xi_n)) - f_n \leq c \alpha_n \nabla f_n^\top$. The choice of the Armijo point becomes important, however, when one also aims for **(A4)** and the convergence rate estimate in Theorem 2.2. To proceed in this direction, we were not able to avoid imposing additional regularity assumptions on the retraction in the limit point.

THEOREM 2.9. *Under the assumptions of Corollary 2.8, assume it holds $\alpha_n \xi_n \rightarrow 0$, and that there exists a constant $C > 0$ such that*

$$\limsup_{n \rightarrow \infty} \frac{\|R(x_n + \hat{\xi}_n) - (x_n + \hat{\xi}_n)\|}{\|\hat{\xi}_n\|^2} \leq C \quad \text{for all sequences } (\hat{\xi}_n) \text{ with } \hat{\xi}_n \in T_n \mathcal{M} \text{ and } \hat{\xi}_n \rightarrow 0. \quad (2.17)$$

Assume further that f is bounded below on whole of \mathbb{R}^N , and that there exists an open (in \mathbb{R}^N) neighborhood \mathcal{N} of x^* such that

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \text{for all } x, y \in \mathcal{N}. \quad (2.18)$$

Then **(A4)** holds (with a generally unknown constant κ). Consequently, $g_n^- \rightarrow 0$, and the convergence rate estimates in Theorem 2.2 apply.

Proof. We can assume $g_n^- > 0$ for all n , since otherwise the sequence becomes stationary. Then $\|\alpha_n \xi_n\| > 0$ for all n . We have to show that $\liminf_{n \rightarrow \infty} \|x_{n+1} - x_n\| / g_n^- > 0$. We do this by showing that the assumption $\liminf_{n \rightarrow \infty} \|x_{n+1} - x_n\| / g_n^- = 0$ leads to a contradiction. In the following we consider a subsequence which converges to the limes inferior, but for notational convenience we assume that

$$\lim_{n \rightarrow \infty} \frac{\|x_{n+1} - x_n\|}{g_n^-} = 0. \quad (2.19)$$

Let $m \in (0, 1)$. The fact that $\alpha_n \xi_n \rightarrow 0$ in combination with (2.17) implies

$$m\|\alpha_n \xi_n\| \leq \|x_{n+1} - x_n\|. \quad (2.20)$$

for large enough n . Clearly, $x_n \in \mathcal{N}$ for every n by the descent property of the algorithm. Since $x_n \rightarrow x^* \in \mathcal{N}$ and $\alpha_n \xi_n \rightarrow 0$, we will also have

$$x_n + z \in \mathcal{N} \quad \text{for all } z \text{ with } \|z\| \leq M\beta^{-1}\|\alpha_n \xi_n\| \quad (2.21)$$

when only n is large enough. For simplicity, we hence can assume that (2.20) and (2.21) hold for all n . We distinguish the iterates by two disjoint cases: $\bar{\beta}_n \leq \bar{\alpha}_n$, and $\bar{\beta}_n > \bar{\alpha}_n$. In the first case, we have $\alpha_n = \bar{\beta}_n$ (Proposition 2.7), which by the choice of $\bar{\beta}_n$ in the algorithm according to (2.16) gives

$$\|x_{n+1} - x_n\| \geq m\|\alpha_n \xi_n\| \geq m g_n^-.$$

Assumption (2.19) implies that this happens only a finite number of times. Let us therefore assume that always the second case $\bar{\beta}_n > \bar{\alpha}_n$ enters. In this case, we have $\bar{\alpha}_n \leq \beta^{-1}\alpha_n$ by Proposition 2.7. Hence, by (2.20) and (2.19),

$$\lim_{n \rightarrow \infty} \frac{\|\bar{\alpha}_n \xi_n\|}{g_n^-} \leq \lim_{n \rightarrow \infty} \frac{m^{-1}\beta^{-1}\|x_{n+1} - x_n\|}{g_n^-} = 0. \quad (2.22)$$

We now mimic the arguments that would prove the existence of step-sizes satisfying the strong Wolfe conditions in linear spaces, e.g [33, Lemma 3.1]. For convenience, let $\hat{R}(\alpha) = R(x_n + \alpha \xi_n)$. By the mean value theorem, there exists $\vartheta \in (0, 1)$ such that for $z = \vartheta(\hat{R}(\bar{\alpha}_n) - x_n)$ it holds

$$(\hat{R}(\bar{\alpha}_n) - x_n)^\top \nabla f(x_n + z) = f(\hat{R}(\bar{\alpha}_n)) - f_n = c\bar{\alpha}_n \xi_n^\top \nabla f_n, \quad (2.23)$$

where the second equality follows from the definition (2.15) of $\bar{\alpha}_n$. By (2.12), $\|z\| \leq M\|\bar{\alpha}_n \xi_n\| \leq M\beta^{-1}\|\alpha_n \xi_n\|$ so that $x_n + z \in \mathcal{N}$ by (2.21). Using (2.18), Cauchy-Schwarz, the definition of z , the reverse triangle inequality, and the angle condition (2.14), we can estimate:

$$\begin{aligned} \|z\| \|\hat{R}(\bar{\alpha}_n) - x_n\| &\geq L^{-1} \|\nabla f_n - \nabla f(x_n + z)\| \|\hat{R}(\bar{\alpha}_n) - x_n\| \\ &\geq L^{-1} |\nabla f_n^\top (\hat{R}(\bar{\alpha}_n) - x_n) - c\bar{\alpha}_n \nabla f_n^\top \xi_n| \\ &\geq L^{-1} ((1-c)\omega g_n^- \|\bar{\alpha}_n \xi_n\| - |\nabla f_n^\top (\hat{R}(\bar{\alpha}_n) - (x_n + \bar{\alpha}_n \xi_n))|). \end{aligned}$$

Since we have $\|\bar{\alpha}_n \xi_n\| \geq M^{-1} \|\hat{R}(\bar{\alpha}_n) - x_n\| \geq M^{-1} \|z\|$, we arrive at

$$\frac{\|\bar{\alpha}_n \xi_n\|}{g_n^-} \geq M^{-2} L^{-1} \left((1-c)\omega - \frac{|\nabla f_n^\top (\hat{R}(\bar{\alpha}_n) - (x_n + \bar{\alpha}_n \xi_n))|}{g_n^- \|\bar{\alpha}_n \xi_n\|} \right). \quad (2.24)$$

By assumption, $\|\bar{\alpha}_n \xi_n\| \leq \beta^{-1}\|\alpha_n \xi_n\| \rightarrow 0$. Since ∇f is continuous, it follows from Cauchy-Schwarz, (2.17), and (2.22), that

$$\lim_{k \rightarrow \infty} \frac{|\nabla f_n^\top (\hat{R}(\bar{\alpha}_n) - (x_n + \bar{\alpha}_n \xi_n))|}{g_n^- \|\bar{\alpha}_n \xi_n\|} \leq \lim_{k \rightarrow \infty} \frac{\|\nabla f(x^*)\| C \|\bar{\alpha}_n \xi_n\|^2}{g_n^- \|\bar{\alpha}_n \xi_n\|} = 0.$$

Consequently, passing to the limit in (2.24) yields a contradiction to (2.22). \square

Property (2.17) in Theorem 2.17 holds for instance for smooth retractions on smooth manifold. In particular, when Euclidean projection is used, $x_{n+1} - x_n \rightarrow 0$ in combination with $x_n \rightarrow x^*$ automatically implies $\alpha_n \xi_n \rightarrow 0$ in the smooth case. These facts lead to the following powerful corollary, which would also hold, if, for instance, orthographic projection from the tangent plane to the manifold [3] can be used (which might be only locally possible).

COROLLARY 2.10. *Let f be real-analytic and bounded below. Assume a simple projection (2.13) is used as retraction in Algorithm 1. Further assume a cluster point x^* for the sequence (x_n) produced by Algorithm 1 exists, and that there is an open neighborhood \mathcal{O} of x^* such that $\mathcal{M} \cap \mathcal{O}$ is a real-analytic submanifold of \mathbb{R}^N . Then (A1)–(A4) hold. Consequently, $\lim_{n \rightarrow \infty} x_n = x^*$ with the convergence rate estimates indicated in Theorem 2.2 apply, and $\lim_{n \rightarrow \infty} g_n^- = g^-(x^*) = 0$.*

Proof. By Proposition 2.1 such a cluster point satisfies the Lojasiewicz inequality (A3). Therefore, by Corollary 2.8, $x_n \rightarrow x^*$ and $\lim_{n \rightarrow \infty} g_n^- \rightarrow g^-(x^*)$ (since on a smooth manifold g^- is a smooth function). For completeness, we now sketch the more or less elementary arguments that $\alpha_n \xi_n \rightarrow 0$, and (2.17) hold. Then Theorem 2.9 applies (the Lipschitz condition for the gradient follows from the analyticity assumption). There exists a local diffeomorphism ϕ from a neighborhood of $0 \in T_{x^*}\mathcal{M}$ (which is a linear space now) to \mathcal{M} such that for large enough n we can write $x^* = \phi(0)$, $x_n = \phi(y_n)$, and $T_n\mathcal{M} = \text{ran}(\phi'(y_n))$. The optimality

condition for $x_{n+1} = R(x_n + \alpha_n \xi_n)$ when being the orthogonal projection of $x_n + \alpha_n \xi_n$ is that the error is orthogonal on the tangent space at x_{n+1} , i.e.,

$$0 = \eta^\top \phi'(y_{n+1})^\top (x_{n+1} - (x_n + \alpha_n \xi_n)) \quad \text{for all } \eta \in T_{x^*} \mathcal{M}.$$

As $x_{n+1} - x_n \rightarrow 0$, this implies

$$0 = \lim_{n \rightarrow \infty} \alpha_n \phi'(y_{n+1})^\top \xi_n.$$

Since the smallest singular value of $\phi'(y_{n+1})^\top$ can be uniformly bounded below for n large enough (the limit $\phi'(0)^\top$ has full rank), it follows $\alpha_n \xi_n \rightarrow 0$. Further, for any $\hat{\xi}_n = \phi'(y_n) \hat{\eta}_n$ we have by the best approximation property of R and Taylor's theorem that

$$\|R(x_n + \hat{\xi}_n) - (x_n + \hat{\xi}_n)\| \leq \|\phi(y_n + \hat{\eta}_n) - (\phi(y_n) + \phi'(y_n) \hat{\eta}_n)\| \leq \|\phi''(x_n + \vartheta_n \hat{\xi}_n)\| \|\hat{\eta}_n\|^2$$

for some $\vartheta_n \in (0, 1)$. If $\hat{\xi}_n \rightarrow 0$ for $n \rightarrow \infty$, then it follows

$$\limsup_{n \rightarrow \infty} \frac{\|R(x_n + \hat{\xi}_n) - (x_n + \hat{\xi}_n)\|}{\|\hat{\xi}_n\|^2} \leq \|\phi''(0)\| \|(\phi'(0))^{-1}\|,$$

since ϕ'' is continuous in zero. \square

3. Results for matrix varieties of bounded rank. The space $\mathbb{R}^m \otimes \mathbb{R}^n \cong \mathbb{R}^{m \times n} \cong \mathbb{R}^{mn}$ becomes an Euclidean space when equipped with the Frobenius inner product $\langle X, Y \rangle_F = \text{trace}(X^\top Y)$. The corresponding norm and distance function are denoted by $\|\cdot\|_F$ and dist_F , respectively. Points in that space will now be denoted by X instead of x , tangent vectors by Ξ instead of ξ . Mainly to save space, we prefer in this paper the subspace and tensor product notation over explicit matrix representations. However, if we use the latter (like in the definition of the inner product), then with respect to some fixed orthonormal bases in \mathbb{R}^m and \mathbb{R}^n . For example, writing $X \in \mathcal{U} \otimes \mathcal{V}$ in $\mathbb{R}^m \otimes \mathbb{R}^n$ would mean in $\mathbb{R}^{m \times n}$ that $X = USV^\top$ for some matrices U, S, V with $\text{ran}(U) = \mathcal{U}$ and $\text{ran}(V) = \mathcal{V}$. By Π_S we denote the orthogonal projection onto a subspace \mathcal{S} . Then $(\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}})X$ corresponds to $UU^\top X V V^\top$, where U and V are orthonormal basis representations for \mathcal{U} and \mathcal{V} , respectively.

The aim of this section is to show that the above convergence theory for projected gradient methods is applicable to the real-algebraic variety $\mathcal{M}_{\leq k}$ of matrices with rank at most k (see (1.3)). We consider the problem

$$\min_{X \in \mathcal{M}_{\leq k}} f(X), \tag{3.1}$$

where, as before, $f : \mathbb{R}^{m \times n} \supseteq \mathcal{D} \rightarrow \mathbb{R}$ is continuously differentiable and bounded below. In fact, in the end we will assume that f is real-analytic to ensure the Łojasiewicz inequality.

3.1. Tangent cone and optimality. Here and in the following, we assume that

$$\text{rank}(X) = s \leq k, \quad \mathcal{U} = \text{ran}(X), \quad \mathcal{V} = \text{ran}(X^\top).$$

The following is well-known.

THEOREM 3.1. *The set \mathcal{M}_s of rank- s matrices is a real-analytic submanifold of dimension $(m + n - s)s$. It is dense and relatively open in $\mathcal{M}_{\leq s}$. The tangent space of \mathcal{M}_s at X is*

$$T_X \mathcal{M}_s = (\mathcal{U} \otimes \mathcal{V}) \oplus (\mathcal{U}^\perp \otimes \mathcal{V}) \oplus (\mathcal{U} \otimes \mathcal{V}^\perp). \tag{3.2}$$

The orthogonal projector on $T_X \mathcal{M}_s$ is hence given by

$$\Pi_{T_X \mathcal{M}_s} = \Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}} + \Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}} + \Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}^\perp} = \Pi_{\mathcal{U}} \otimes I + I \otimes \Pi_{\mathcal{V}} - \Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}}, \tag{3.3}$$

and it holds

$$T_X \mathcal{M}_s \oplus (\mathcal{U}^\perp \otimes \mathcal{V}^\perp) = \mathbb{R}^m \otimes \mathbb{R}^n. \quad (3.4)$$

Our main task is to investigate the tangent cones of $\mathcal{M}_{\leq k}$ in points with $s < k$. The tangent cone $T_X \mathcal{M}_{\leq k}$ clearly contains $T_X \mathcal{M}_s$, but, in case $s < k$, also rays that arise when approaching X by a matrix of rank at most k , but larger than s .

THEOREM 3.2. [12] *Let $X \in \mathcal{M}_{\leq k}$, $\text{rank}(X) = s$. The tangent cone of $\mathcal{M}_{\leq k}$ at X is*

$$T_X \mathcal{M}_{\leq k} = T_X \mathcal{M}_s \oplus \{\Xi_{k-s} \in \mathcal{U}^\perp \otimes \mathcal{V}^\perp : \text{rank}(\Xi_{k-s}) \leq k - s\}.$$

Proof. To prove the “ \supseteq ” part, let Ξ be an element from set on the right side of the equality. Then $\Xi = \Xi_s + \Xi_{k-s}$ with $\Xi_s \in T_X \mathcal{M}_s$, and $\text{rank}(\Xi_{k-s}) \leq k - s$. There exists a sequence $(Y_n) \subseteq \mathcal{M}_s$, and a sequence $(a_n) \subseteq \mathbb{R}^+$ such that $Y_n \rightarrow X$, and $a_n(Y_n - X) = \Xi_s$. One can assume $a_n \rightarrow \infty$ (i.e. $\Xi_s \neq 0$). Then $X_n = Y_n + a_n^{-1} \Xi_{k-s}$ is a sequence in $\mathcal{M}_{\leq k}$ which converges to X , and $a_n(X_n - X)$ converges to Ξ , which proves $\Xi \in T_X \mathcal{M}_{\leq k}$.

To prove the reverse inclusion “ \subseteq ”, assume $\Xi = \lim_{n \rightarrow \infty} a_n(X_n - X)$, $X_n \rightarrow X$ in $\mathcal{M}_{\leq k}$, and $(a_n) \subseteq \mathbb{R}^+$. In the orthogonal decomposition

$$a_n(X_n - X) = \Pi_{T_X \mathcal{M}_s} a_n(X_n - X) + (\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}^\perp}) a_n X_n$$

both terms have to converge separately. Denote their limits by Ξ_s and Ξ_{k-s} , respectively. Then obviously $\Xi = \Xi_s + \Xi_{k-s}$ with $\Xi_s \in T_X \mathcal{M}_s$ and $\Xi_{k-s} \in \mathcal{U}^\perp \otimes \mathcal{V}^\perp$. Since $(\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}}) X_n \rightarrow (\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}}) X = X$, and since the set of rank- s matrices is relatively open in $\mathcal{U} \otimes \mathcal{V}$, $\text{rank}((\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}}) X_n) = s$ for large enough n . Consequently, since $\text{rank}(X_n) \leq k$ for all n , it must hold $\text{rank}((\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}^\perp}) a_n X_n) = \text{rank}((\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}^\perp}) X_n) \leq k - s$ for large enough n . It follows by the semi-continuity of matrix rank, that $\text{rank}(\Xi_{k-s}) \leq k - s$. \square

Remark. In the recent paper [12] the tangent cones of $\mathcal{M}_{\leq k}$ have been previously derived, but in contrast to (2.2) have been defined via analytic curves as

$$T_X \mathcal{M}_{\leq k} = \{\dot{\gamma}(0) : \gamma(t) \text{ is an analytic curve with } \gamma(t) \in \mathcal{M}_{\leq k} \text{ for } t \geq 0, \text{ and } \gamma(0) = X\}. \quad (3.5)$$

As shown in [37, Proposition 2] both definitions are equivalent. Up to an additional normalization constraint, the authors of [12] essentially prove Theorem 3.2 using definition (3.5). As an outcome of our proof, we have a direct verification that both definitions are equivalent. As mentioned in [12], when using the definition (3.5), the “ \subseteq ” part in Theorem 3.2 follows from known results on the existence of analytic “singular value decomposition paths” [10]. We can easily modify our argument above to prove the “ \supseteq ” part for (3.5) by choosing an analytic curve γ_s in \mathcal{M}_s such that $\Xi_s = \dot{\gamma}_s(0)$, and put $\gamma(t) = \gamma_s(t) + t \Xi_{k-s}$, which is an analytic curve in $\mathcal{M}_{\leq k}$ with $\dot{\gamma}(0) = \Xi = \Xi_s + \Xi_{k-s}$. The proof of “ \supseteq ” given in [12] seems more involved than probably necessary, since the well-known structure of $T_X \mathcal{M}_s$ is not exploited.

Now that we know the structure of the tangent cone in rank-deficient points, we can calculate the projection of the negative gradient on it. This turns out to be very easy. Moreover, the tangent cone in such points is so large that the projection on it carries over astonishingly much information.

COROLLARY 3.3. *Let $X \in \mathcal{M}_{\leq k}$, $\text{rank}(X) = s$. Any $G \in T_X \mathcal{M}_{\leq k}$ with $\| -\nabla f(X) - G \|_F = \text{dist}_F(-\nabla f(X), T_X \mathcal{M}_{\leq k})$ has the form*

$$G = \Pi_{T_X \mathcal{M}_s}(-\nabla f(X)) + \Xi_{k-s}, \quad (3.6)$$

where Ξ_{k-s} is a best rank- $(k-s)$ approximation of $(\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}^\perp})(-\nabla f(X)) = -\nabla f(X) - \Pi_{T_X \mathcal{M}_s}(-\nabla f(X))$ in the Frobenius norm. (Obviously, $\Xi_{k-s} \in \mathcal{U}^\perp \otimes \mathcal{V}^\perp$ then.) Moreover,

$$g^-(X) = \|G\|_F \geq \sqrt{\frac{k-s}{\min(m,n)}} \|\nabla f(X)\|_F. \quad (3.7)$$

Proof. The form of G is clear from Theorem 3.2 by orthogonality considerations. We prove the norm estimate. The square of the Frobenius norm of a matrix is the sum of its squared singular values, of which there are at most $\min(m, n)$. The best rank- $(k - s)$ approximation of a matrix in the Frobenius norm is obtained by truncating its singular value decomposition up to the largest $k - s$ singular values. We conclude that

$$\|\Xi_{k-s}\|_F^2 \geq \frac{k-s}{\min(m, n)} \|(\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}^\perp}) \nabla f(X)\|_F^2,$$

which with (3.6) and (3.4) already proves the bound. \square

This leads us to a remarkable, at first sight probably unexpected, a-priori statement about critical points of differentiable functions on $\mathcal{M}_{\leq k}$. It seems to be new.

COROLLARY 3.4. *Let $k \leq \min(m, n)$, and let $X^* \in \mathcal{M}_{\leq k}$ be a critical point of (3.1) in the sense $g^-(X^*) = 0$. Then either $\text{rank}(X^*) = k$ or $\nabla f(X^*) = 0$.*

As an illustration consider the following.

COROLLARY 3.5. *Let $k \leq \min(m, n)$. Assume $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ is strictly convex, coercive, and its unique minimizer on $\mathbb{R}^{m \times n}$ has rank larger than or equal to k . Then any relative local minimizer of f on $\mathcal{M}_{\leq k}$ has rank k .*

In light of these results, it is not surprising that we will have to make the assumption $\text{rank}(X^*) = k$ in our convergence results below in order to conclude $g^-(X^*) = 0$. It is not an artifact of the used techniques. Instead, Corollary 3.4 tells us that it will be normally impossible to find a rank-deficient critical point by a projected gradient method that most of the time moves on \mathcal{M}_k , since on \mathcal{M}_k the projection of the gradient contains much less information.

We finish with a practical remark. When the matrices are huge, one will only be able to work with low-rank representations of all involved quantities. In particular, $\nabla f(X)$ needs to allow for a sparse or a low-rank representation. If $s \ll \min(m, n)$, the calculation of $\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}^\perp} (-\nabla f(X)) = -\nabla f(X) - \Pi_{T_X \mathcal{M}_s} (-\nabla f(X))$ is then feasible using the second representation of $\Pi_{T_X \mathcal{M}_s}$ in (3.3). With some effort one can even exploit the low-rank structure of $\Pi_{T_X \mathcal{M}_s} (-\nabla f(X))$ to calculate an SVD of the difference without forming it. The huge projector $\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}^\perp}$ should never be formed. The final rank of tangent vectors itself is not larger than $2s + (k - s) = k + s$, which can be seen from the decomposition. We summarize the procedure as Algorithm 2.

Algorithm 2: Calculate the projection of $-\nabla f(X)$ on $T_X \mathcal{M}_{\leq k}$

Input: Negative gradient $F = -\nabla f(X)$ at $X \in \mathcal{M}_{\leq k}$.

Output: Projection $G \in T_X \mathcal{M}_{\leq k}$ with $\|F - G\|_F = \text{dist}_F(F, T_X \mathcal{M}_{\leq k})$

- 1 Find orthonormal bases U and V for $\text{ran}(X)$ and $\text{ran}(X^\top)$, respectively;
- 2 Calculate the projection on $T_X \mathcal{M}_s$:

$$\Xi_s = UU^\top F + FVV^\top - UU^\top FVV^\top;$$

- 3 Calculate best rank- $(k - s)$ approximation of the difference:

$$\Xi_{k-s} \in \underset{\text{rank}(Y) \leq k-s}{\text{argmin}} \|F - \Xi_s - Y\|_F;$$

- 4 Output:

$$G = \Xi_s + \Xi_{k-s}, \quad g^-(X) = \|G\|_F = \sqrt{\|\Xi_s\|_F^2 + \|\Xi_{k-s}\|_F^2}.$$

3.2. Retraction by best low-rank approximation. As a retraction we choose the best approximation by a matrix of rank at most k in the Frobenius norm, i.e.,

$$R(X + \Xi_X) \in \operatorname{argmin}_{Y \in \mathcal{M}_{\leq k}} \|Y - (X + \Xi_X)\|_F. \quad (3.8)$$

It can be explicitly calculated using singular value decomposition. In unlikely events, (3.8) is set-valued, but we can assume that a specific choice is made by fixing a deterministic SVD and sorting algorithm for the singular values. The particular choice does not matter. Let us emphasize once more that Definition 2.3 is indeed fulfilled: Let $\Xi \in T_X \mathcal{M}_{\leq k}$, then by [37, Proposition 2] there exists an analytic arc $\gamma : [0, \epsilon) \rightarrow \mathcal{M}_{\leq k}$ such that $\dot{\gamma}(0) = \Xi$. Hence,

$$\lim_{\alpha \rightarrow 0^+} \frac{\|R(X + \alpha \Xi_X) - (X + \Xi_X)\|_F}{\alpha} \leq \lim_{\alpha \rightarrow 0^+} \frac{\|\gamma(\alpha) - (X + \dot{\gamma}(0))\|_F}{\alpha} = 0. \quad (3.9)$$

We have the following nice estimate, which provides $M = 1 + 2^{-1/2}$ in (2.12).

PROPOSITION 3.6. *The above retraction satisfies*

$$\|R(X + \Xi_X) - (X + \Xi_X)\|_F \leq \frac{1}{\sqrt{2}} \|\Xi_X\|_F \quad \text{for all } X \in \mathcal{M}_{\leq k} \text{ and } \Xi_X \in T_X \mathcal{M}_{\leq k}.$$

Proof. The matrices $X + (\Pi_U \otimes I)\Xi_X = (\Pi_U \otimes I)(X + \Xi_X)$, and $X + (I \otimes \Pi_V)\Xi_X = (I \otimes \Pi_V)(X + \Xi_X)$ both have rank at most s . Thus, by Theorem 3.2,

$$X + (\Pi_U \otimes I)\Xi_X + (\Pi_{U^\perp} \otimes \Pi_{V^\perp})\Xi_X,$$

and

$$X + (I \otimes \Pi_V)\Xi_X + (\Pi_{U^\perp} \otimes \Pi_{V^\perp})\Xi_X,$$

both have rank not larger than k . Considering them as possible candidates for a best approximation $R(X + \Xi_X)$ of $X + \Xi_X$ by a matrix of rank at most k , we obtain the desired bound

$$\|R(X + \Xi_X) - (X + \Xi_X)\|_F^2 \leq \min(\|(\Pi_{U^\perp} \otimes \Pi_V)\Xi_X\|_F^2, \|(\Pi_U \otimes \Pi_{V^\perp})\Xi_X\|_F^2) \leq \frac{1}{2} \|\Xi_X\|_F^2,$$

where we have made use of the orthogonal decompositions (3.2) and (3.4). \square

As a consequence, we observe that $\|R(X + \Xi_X) - X\|_F^2 \rightarrow 0$ automatically implies $\Xi_X \rightarrow 0$.

3.3. Łojasiewicz inequality and convergence result. To apply the convergence results of section 2.3, we show that the Łojasiewicz inequality (A3) holds in every point of $\mathcal{M}_{\leq k}$, thereby considering the simplest case of real-analytic functions. Somewhat surprisingly, the proof of the inequality in the singular points of $\mathcal{M}_{\leq k}$ is in a way even simpler than in regular points, since it directly follows from the Łojasiewicz inequality in linear spaces. Therefore it should pose no difficulties to obtain more general results for larger classes of functions.

THEOREM 3.7. *Let $\mathcal{D} \subseteq \mathbb{R}^{m \times n}$ be open, $\mathcal{M}_{\leq k} \subset \mathcal{D}$, and $f : \mathcal{D} \rightarrow \mathbb{R}$ be real-analytic. Then the Łojasiewicz inequality (A3) holds at any point $X \in \mathcal{M}_{\leq k}$.*

Proof. For regular points X with $\operatorname{rank}(X) = k$, this follows from Theorem 3.1 and Proposition 2.1. For X with $\operatorname{rank}(X) = s < k$ we apply Proposition 2.1 to $\mathcal{M} = \mathcal{D}$, and use (3.7). \square

We have now collected all requirements to apply Theorem 2.9 or Corollary 2.10. For concreteness, we consider a particular algorithm where the search direction is the projected negative gradient, and the retraction is obtained by best rank- k approximation. It is denoted as Algorithm 3.

THEOREM 3.8. *Let f be real-analytic and bounded below. If the sequence (X_n) generated by Algorithm 3 possesses a cluster point X^* , then it is its limit, $\lim_{n \rightarrow \infty} X_n = X^*$. If further $\operatorname{rank}(X^*) = k$, then $g^-(X^*) = 0$, and the convergence rate estimates of Theorem 2.2 apply.*

Proof. The convergence of the sequence follows from Theorem 3.7, Proposition 2.5, and Corollary 2.8. Due to Theorem 3.1, the rest is an instance of Corollary 2.10. \square

Algorithm 3: Projected steepest descent with line-search on $\mathcal{M}_{\leq k}$

Input: Starting guess $X_0 \in \mathcal{M}_{\leq k}$, $\beta, c \in (0, 1)$.

- 1 **for** $n=0, 1, 2, \dots$ **do**
- 2 Calculate a projection G_n of $-\nabla f(X_n)$ on $T_{X_n}\mathcal{M}_{\leq k}$ using Algorithm 2;
- 3 Choose $\bar{\beta}_n \geq 1$, and find Armijo point α_n for $X_n, G_n, \bar{\beta}_n, \beta, c$;
- 4 Set X_{n+1} to be a best approximation (w.r.t Frobenius norm) of $X_n + \alpha_n G_n$ of rank at most k .
- 5 **end**

3.4. A method without retraction. It is possible to have a gradient-related search direction Ξ_n such that $X_n + \alpha \Xi_n \in \mathcal{M}_{\leq k}$ for all α . The idea is the same as in the proof of Proposition 3.6. By (3.6) and (3.2), a projection G of $-\nabla f_n$ consists of four, mutually orthogonal parts:

$$G_n = (\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}})(-\nabla f_n) + (\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}})(-\nabla f_n) + (\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}^\perp})(-\nabla f_n) + \Xi_{k-s,n},$$

with $\text{rank}(\Xi_{k-s,n}) \leq k - s$. Consider the two possible partial projections

$$G_n^{(1)} = (\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}})(-\nabla f_n) + (\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}^\perp})(-\nabla f_n) + \Xi_{k-s,n} = (\Pi_{\mathcal{U}} \otimes I)(-\nabla f_n) + \Xi_{k-s,n}, \quad (3.10)$$

and

$$G_n^{(2)} = (\Pi_{\mathcal{U}} \otimes \Pi_{\mathcal{V}})(-\nabla f_n) + (\Pi_{\mathcal{U}^\perp} \otimes \Pi_{\mathcal{V}})(-\nabla f_n) + \Xi_{k-s,n} = (I \otimes \Pi_{\mathcal{V}})(-\nabla f_n) + \Xi_{k-s,n}. \quad (3.11)$$

Both are elements of the tangent cone at X_n and satisfy $\text{rank}(X_n + \alpha G_n^{(i)}) \leq k$ for all α , $i = 1, 2$. Assume that $\|G_n^{(1)}\|_{\text{F}} \geq \|G_n^{(2)}\|_{\text{F}}$. Then, by orthogonality arguments, $\|G_n^{(1)}\|_{\text{F}}^2 \geq \frac{1}{2}\|G_n^{(2)}\|_{\text{F}}^2$, and

$$\langle \nabla f_n, G_n^{(1)} \rangle_{\text{F}} = \|G_n^{(1)}\|_{\text{F}}^2 \geq \frac{1}{\sqrt{2}}\|G_n\|_{\text{F}}\|G_n^{(1)}\|_{\text{F}} = \frac{1}{\sqrt{2}}g_n^-\|G_n^{(1)}\|_{\text{F}}.$$

Thus $G_n^{(1)}$ satisfies the angle condition with $\omega = \frac{1}{\sqrt{2}}$. If $\|G_n^{(1)}\|_{\text{F}} \leq \|G_n^{(2)}\|_{\text{F}}$, then $G_n^{(2)}$ satisfies this angle condition.

This leads us to Algorithm 4, which contains no retraction steps. Still, it shares the nice abstract convergence features as the classical steepest descent, even with a slightly extended statement in singular points.

THEOREM 3.9. *Let f be real-analytic and bounded below. If the sequence (X_n) generated by Algorithm 4 possesses a cluster point X^* , then it is its limit, $\lim_{n \rightarrow \infty} X_n = X^*$, and the convergence rate estimates of Theorem 2.2 apply. If further $\text{rank}(X^*) = k$, then $g^-(X^*) = 0$.*

Proof. Since $\text{rank}(X_n + \alpha_n \Xi_n) \leq k$, we can formally write $X_{n+1} = R(X_n + \alpha_n \Xi_n)$ in the algorithm in order to get into the abstract framework (here R is the retraction by best low-rank approximation). Then the mere convergence of the sequence follows again from Theorem 3.7 and Corollary 2.8. The feature is now that (2.17) is trivially satisfied, therefore the validity of the convergence rate estimates follows from Theorem 2.9 even in singular limit points (the Lipschitz condition (2.18) follows from analyticity). We also have $g_n^-(X_n) \rightarrow 0$ from which we can conclude $g^-(X^*) = 0$ if g^- is continuous in X^* , which in regular points $X^* \in \mathcal{M}_k$ is the case. \square

As it does not leave the feasible set, Algorithm 4 is very elegant and saves some cost in every step of the back-tracking to find the Armijo point. In applications, however, the retraction from rank (at most) $2k$ to rank k , as required in Algorithm 3, is typically much less expensive than for instance a function value evaluation or the projection of the gradient. We hence expect that the saved retractions will seldomly compensate for the less gradient-related search directions of Algorithm 4.

Algorithm 4: Descent method on $\mathcal{M}_{\leq k}$ without retraction

Input: Starting guess $X_0 \in \mathcal{M}_{\leq k}$, $\beta, c \in (0, 1)$.

```

1 for  $n=0, 1, 2, \dots$  do
2   if  $\|(\Pi_{\mathcal{U}} \otimes I)(-\nabla f(X_n))\|_{\mathbb{F}} \geq \|(I \otimes \Pi_{\mathcal{V}})(-\nabla f(X_n))\|_{\mathbb{F}}$  then
3     | Use  $\Xi_n = G_n^{(1)}$  from (3.10);
4   else
5     | Use  $\Xi_n = G_n^{(2)}$  from (3.11);
6   end
7   Choose  $\bar{\beta}_n \geq \sqrt{2}$ , and find Armijo point  $\alpha_n$  for  $X_n, \Xi_n, \bar{\beta}_n, \beta, c$ ;
8   Form the next iterate
      
$$X_{n+1} = X_n + \alpha_n \Xi_n.$$

9 end

```

We checked this with a toy example of matrix completion in a setup similar to [42], using *straight-forward, comparably non-optimized* MATLAB R2012b implementations of both algorithms (choosing $\beta = \frac{1}{2}$ and $c = 10^{-4}$) on a Notebook with a 1.9 GHz CPU and 4 GB of memory. The problem that was solved is

$$\min_{X \in \mathcal{M}_{\leq k}} \frac{1}{2} \|P_{\Omega}(A - X)\|_{\mathbb{F}}^2, \quad (3.12)$$

where P_{Ω} is the projector on a subset Ω of indices. The $n \times n$ matrix A of rank (at most) k was randomly generated by randomly generating two $n \times k$ factor matrices from a normal distribution. Hence the global minimum of (3.12) is $X^* = A$. The starting guess X_0 (used for both algorithms) was created in the same way. The size of Ω was chosen as $|\Omega| = \max(\text{OS} \cdot (2kn - k^2), n \log n)$, which is an oversampling rate of at least OS (cf. [42]), and Ω itself was drawn uniformly at random. For $n = 2000$, $k = 20$, and $\text{OS} = 3$ this means 94, 03% missing entries. For this case, the relative errors,

$$\frac{\|A - X_n\|_{\mathbb{F}}}{\|A\|_{\mathbb{F}}}, \quad (3.13)$$

as well as the relative errors on the visible index set,

$$\frac{\|P_{\Omega}(A - X_n)\|_{\mathbb{F}}}{\|P_{\Omega}A\|_{\mathbb{F}}} = \frac{\sqrt{2f(X_n)}}{\|P_{\Omega}A\|_{\mathbb{F}}}, \quad (3.14)$$

are plotted in Figure 9. As one can see, Algorithm 4 is inferior to Algorithm 3 with respect to both, number of iterations and computation time (the latter is given to have an impression). One might think that the relative performance of Algorithm 4 improves for larger k . The plots for $k = 80$ do not support this hope (in this case only 76, 48% entries are missing, which explains the faster error decay).

Both algorithms served only as examples, and are naturally inferior to more sophisticated line-search methods, like the nonlinear CG methods used in [42], which use gradient information from previous iterates.

4. Conclusion. We extended previous results on convergence of descent iterations on manifolds using Lojasiewicz inequality to gradient-related projection methods on the real-algebraic variety $\mathcal{M}_{\leq k}$ of real $m \times n$ matrices of rank at most k . So far, the results are applicable when the cost function is real-analytic, as it is the case for many applications. It is our hope that in particular the studies on the tangent cones of $\mathcal{M}_{\leq k}$ at singular points will help to understand and circumvent the problems of non-closedness and

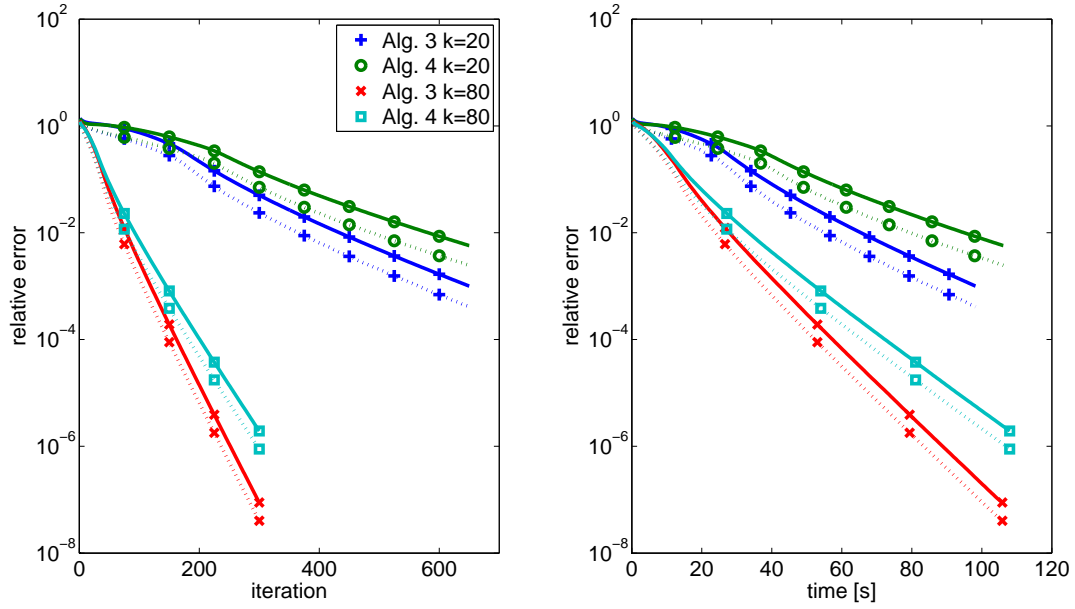


FIG. 3.1. Comparison of Algorithms 3 and 4 for the matrix completion problem (3.12) with $A \in \mathbb{R}^{2000 \times 2000}$, $k = 20$ (94, 03% missing entries), and rank $k = 80$ (76, 48% missing entries). Solid lines: relative errors (3.13) (3.13) (full index set); dashed lines: relative error (3.14) (sample index set).

unbounded curvature, that occur in the convergence analysis of Riemannian optimization methods on the smooth manifold \mathcal{M}_k of fixed rank matrices, in the future.

This hope also comprises manifolds of low-rank tensors, where similar issues are encountered. There is growing interest in treating low-rank tensor problems by Riemannian optimization, e.g. tensor completion [20], or dynamical tensor approximation [18, 28, 41]. It would hence be important and interesting to extend the results to tensor varieties of bounded subspace ranks, e.g. bounded Tucker ranks, hierarchical Tucker ranks or tensor train ranks [19, 14, 36]. A first idea based on the results of this paper could be that these varieties take the form of intersections of certain low-rank matrix varieties [41].

REFERENCES

- [1] P.-A. ABSIL, R. MAHONY, AND B. ANDREWS, *Convergence of the iterates of descent methods for analytic cost functions*, SIAM J. Optim., 16 (2005), pp. 531–547.
- [2] P.-A. ABSIL, R. MAHONY, AND R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [3] P.-A. ABSIL AND J. MALICK, *Projection-like retractions on matrix manifolds*, SIAM J. Optim., 22 (2012), pp. 135–158.
- [4] H. ATTOUCH AND J. BOLTE, *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features*, Math. Program., 116 (2009), pp. 5–16.
- [5] H. ATTOUCH, J. BOLTE, P. REDONT, AND A. SOUBEYRAN, *Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality*, Math. Oper. Res., 35 (2010), pp. 438–457.
- [6] H. ATTOUCH, J. BOLTE, AND B. F. SVAITER, *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods*, Math. Program., 137 (2013), pp. 91–129.
- [7] J. BOLTE, A. DANIILIDIS, AND A. LEWIS, *The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems*, SIAM J. Optim., 17 (2006), pp. 1205–1223.
- [8] J. BOLTE, A. DANIILIDIS, A. LEWIS, AND M. SHIOTA, *Clarke subgradients of stratifiable functions*, SIAM J. Optim., 18

- (2007), pp. 556–572.
- [9] J. BOLTE, A. DANIILIDIS, O. LEY, AND L. MAZET, *Characterizations of tojasiewicz inequalities: subgradient flows, talweg, convexity*, Trans. Amer. Math. Soc., 362 (2010), pp. 3319–3363.
 - [10] A. BUNSE-GERSTNER, R. BYERS, V. MEHRMANN, AND N. K. NICHOLS, *Numerical computation of an analytic singular value decomposition of a matrix valued function*, Numer. Math., 60 (1991), pp. 1–39.
 - [11] E. CANCÈS, V. EHRLACHER, AND T. LELIÈVRE, *Greedy algorithms for high-dimensional eigenvalue problems*, arXiv preprint arXiv:1304.2631, (2013).
 - [12] T. P. CASON, P.-A. ABSIL, AND P. VAN DOOREN, *Iterative methods for low rank approximation of graph similarity matrices*, Linear Algebra Appl., 438 (2013), pp. 1863–1882.
 - [13] M. GUIGNARD, *Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space*, SIAM J. Control, 7 (1969), pp. 232–241.
 - [14] W. HACKBUSCH AND S. KÜHN, *A new scheme for the tensor representation*, J. Fourier Anal. Appl., 15 (2009), pp. 706–722.
 - [15] A. HARAUX AND M. A. JENDOUBI, *The Łojasiewicz gradient inequality in the infinite-dimensional Hilbert space framework*, J. Funct. Anal., 260 (2011), pp. 2826–2842.
 - [16] S.-Z. HUANG, *Gradient Inequalities. With applications to asymptotic behavior and stability of gradient-like systems*, American Mathematical Society, Providence, RI, 2006.
 - [17] O. KOCH AND C. LUBICH, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 434–454.
 - [18] ———, *Dynamical tensor approximation*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2360–2375.
 - [19] T. G. KOLDA AND B. W. BADER, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.
 - [20] D. KRESSNER, M. STEINLECHNER, AND B. VANDEREYCKEN, *Low-rank tensor completion by Riemannian optimization*, BIT, (2013). In press.
 - [21] K. KURDYKA, *On gradients of functions definable in o-minimal structures*, Ann. Inst. Fourier (Grenoble), 48 (1998), pp. 769–783.
 - [22] C. LAGEMAN, *Konvergenz reell-analytischer gradientenähnlicher Systeme*, diploma thesis, Universität Würzburg, Würzburg, Germany. In German.
 - [23] ———, *Convergence of gradient-like dynamical systems and optimization algorithms*, PhD thesis, Universität Würzburg, Würzburg, Germany, 2007.
 - [24] ———, *Pointwise convergence of gradient-like systems*, Math. Nachr., 280 (2007), pp. 1543–1558.
 - [25] A. LEVITT, *Convergence of gradient-based algorithms for the Hartree-Fock equations*, ESAIM Math. Model. Numer. Anal., 46 (2012), pp. 1321–1336.
 - [26] Z. LI, A. USCHMAJEW, AND S. ZHANG, *On convergence of the maximum block improvement method*, MATHICSE Report 30.2013, EPF Lausanne, 2013.
 - [27] S. ŁOJASIEWICZ, *Ensemble semi-analytique*. Note des cours, Institut des Hautes Etudes Scientifique, 1965.
 - [28] C. LUBICH, T. ROHWEDDER, R. SCHNEIDER, AND B. VANDEREYCKEN, *Dynamical approximation by hierarchical Tucker and tensor-train tensors*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 470–494.
 - [29] C. LUBICH AND I. V. OSELEDTS, *A projector-splitting integrator for dynamical low-rank approximation*, BIT, (2013). In press.
 - [30] B. MERLET AND T. N. NGUYEN, *Convergence to equilibrium for discretizations of gradient-like flows on Riemannian manifolds*, Differential Integral Equations, 26 (2013), pp. 571–602.
 - [31] B. MISHRA, G. MEYER, F. BACH, AND R. SEPULCHRE, *Low-rank optimization with trace norm penalty*, SIAM J. Optim., 23 (2013), pp. 2124–2149.
 - [32] B. MISHRA, G. MEYER, S. BONNABEL, AND R. SEPULCHRE, *Fixed-rank matrix factorizations and Riemannian low-rank optimization*, Comput. Statist., (2013). In press.
 - [33] J. NOCEDAL AND S. J. WRIGHT, *Numerical Optimization*, Springer, New York, 2006.
 - [34] D. NOLL, *Convergence of non-smooth descent methods using the Kurdyka–Łojasiewicz inequality*, J. Optim. Theory Appl., (2013). In press.
 - [35] A. NONNENMACHER AND C. LUBICH, *Dynamical low-rank approximation: applications and numerical experiments*, Math. Comput. Simulation, 79 (2008), pp. 1346–1357.
 - [36] I. V. OSELEDTS, *Tensor-train decomposition*, SIAM J. Sci. Comput., 33 (2011), pp. 2295–2317.
 - [37] D. B. O’SHEA AND L. C. WILSON, *Limits of tangent spaces to real surfaces*, Amer. J. Math., 126 (2004), pp. 951–980.
 - [38] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational Analysis*, Springer-Verlag, Berlin, 1998.
 - [39] U. SHALIT, D. WEINSHALL, AND G. CHECHIK, *Online learning in the embedded manifold of low-rank matrices*, J. Mach. Learn. Res., 13 (2012), pp. 429–458.
 - [40] M. SHUB, *Some remarks on dynamical systems and numerical analysis*, in Dynamical systems and partial differential equations (Caracas, 1984), Univ. Simon Bolivar, Caracas, 1986, pp. 69–91.
 - [41] A. USCHMAJEW AND B. VANDEREYCKEN, *The geometry of algorithms using hierarchical tensors*, Linear Algebra Appl., 439 (2013), pp. 133–166.
 - [42] B. VANDEREYCKEN, *Low-rank matrix completion by Riemannian optimization*, SIAM J. Optim., 23 (2013), pp. 1214–1236.
 - [43] B. VANDEREYCKEN AND S. VANDEWALLE, *A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2553–2579.
 - [44] Y. XU AND W. YIN, *A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion*, SIAM J. Imaging Sci., 6 (2013), pp. 1758–1789.